TECHNICAL MEMORANDUM

Sonoma Technology, Inc.

1360 Redwood Way, Suite C
Petaluma, CA 94954
707/665-9900
FAX 707/665-9800
www.sonomatech.com

August 26, 2002 STI Ref. No. 901491

TO: Ms. Cheryl Heying, Utah Department of Environmental Quality

COPY: Mr. John White, U.S. Environmental Protection Agency

FROM: Timothy Dye, Dianne Miller, and Clinton MacDonald

SUBJECT: Summary of $PM_{2.5}$ forecasting program development and operations for Salt Lake City, Utah during winter 2002.

Sonoma Technology, Inc. (STI) has completed the "Pilot PM Forecasting" project for Salt Lake City. This letter report describes the goals of the project, major task activities, forecasting methods, performance results of the 45-day forecasting period, and recommendations for future work. It is intended to serve as a technical transfer of information from STI to Utah Department of Environmental Quality (UDEQ) and provide the basis for UDEQ to further develop their local PM forecasting program. Additional information on developing an air quality forecasting program can be found at U.S. Environmental Protection Agency (1999). In addition, our staff is available to answer questions about this project.

We appreciate all of the help from UDEQ staff during this pilot project and look forward to opportunities to work with UDEQ in the future.

Sincerely,

Timothy S. Dye
Senior Manager, Meteorological Programs Group

1. INTRODUCTION

Atmospheric particulate matter less than 2.5 μm in size ($PM_{2.5}$) is a complex mixture of chemical constituents resulting from primary emissions of fine particles and emissions of gaseous compounds (SO_2, NO_x, VOC, and NH_3) that undergo chemical reactions to form secondary aerosols. Many different types of anthropogenic sources contribute to ambient concentrations of $PM_{2.5}$. Studies indicate that short-term exposure to acute $PM_{2.5}$ concentrations can lead to long-term health effects. The negative health effects associated with high $PM_{2.5}$ concentrations and the public's desire for accurate air quality information has produced a need for $PM_{2.5}$ forecasting programs that warn the public one or two days in advance of high $PM_{2.5}$ concentrations.

The U.S. Environmental Protection Agency's (EPA) AIRNow forecasting program is currently focused on summertime air pollutants (such as ozone), so the program does not capture year-round air pollutants like PM. By helping to develop local $PM_{2.5}$ forecasting programs, EPA seeks to provide a consistent, year-round air quality forecast for the public and for media outlets such as *USA Today* and The Weather Channel. To achieve this expansion, EPA funded a pilot $PM_{2.5}$ forecasting project for two cities during 2002. The major goal of this project was to develop and implement $PM_{2.5}$ forecasting techniques and then to conduct pilot $PM_{2.5}$ forecasting studies for Salt Lake City, Utah and Pittsburgh, Pennsylvania.

Participants in this pilot program included the EPA's AIRNow program staff, who funded the project and used the daily forecasts; Utah Department of Environmental Quality (UDEQ) staff, who provided data, local support, and used the daily forecasts; and staff at Sonoma Technology, Inc. (STI), who developed the forecasting methods and issued the daily forecasts. Forecasts were issued for a 45-day period (January 15 to February 28, 2002), which coincided with the 2002 Winter Olympic Games in Salt Lake City. Details about the project, forecasting methods, forecast performance, and recommendations for future work are provided in the remainder of this report.

2. PROJECT ACTIVITIES

This section provides an overview of the major tasks that STI completed during this project. Some technical details are provided, but more detailed technical information about air quality forecasting can be found in (Dye et al., 2002; U.S. Environmental Protection Agency, 1999).

1. <u>Coordination</u>. STI worked closely with staff at UDEQ to obtain access to real-time $PM_{2.5}$ data via the Internet. In addition, we worked to educate UDEQ staff about all aspects of this project and attended an onsite meeting to answer questions, obtain historical data, and develop an understanding of the meteorology in the Salt Lake City region.

2. <u>Data Acquisition and Database Development</u>. Forecasting program development requires access to historical meteorological and air quality data to understand how weather affects pollution levels. Once a program is developed, access to real-time weather and air

quality data are needed to help formulate the forecast. For this project, STI acquired the following historical data for the winter months (October through March):

- Five years of PM_{10} data (1996 – 2001) from 14 sites
- Two to four years of $PM_{2.5}$ data (1998 – 2001) from 16 sites
- Continuous $PM_{2.5}$ and PM_{10} data from three Tapered Element Oscillating Microbalance (TEOM) monitors (1999 – 2001)
- Five years of hourly surface observations and upper-air soundings from Salt Lake City airport
- Daily weather maps

Data were decoded, assigned common units and time standards, and stored in a Microsoft Access database. STI staff then reviewed the PM and meteorological data for quality problems. Visual inspections of data and statistical measures (maximum, minimum, standard deviation) were reviewed to identify outliers in the data, which were flagged and excluded from the data analysis.

3. Data analysis and method development. Several steps were performed to develop the forecasting tools. After collecting all the historical data, a PM climatology was developed to study long-term trends. Next, a conceptual model was developed to identify and understand how meteorological processes affect PM; this conceptual understanding helped develop forecasting guidelines and select variables for use in regression equations. Finally, statistical analysis was performed using scatter plots, correlations, and factor analysis to determine the final set of variables to be used for developing statistical forecasting methods.

   The air quality climatology helped the meteorologists better understand the nature of the characteristics of PM episodes in the region. Annual, monthly, weekly, and day-of-week frequencies of Air Quality Index (AQI) categories (based on 24-hr average $PM_{2.5}$ concentrations) were analyzed for the entire Salt Lake City basin. In addition, diurnal patterns of $PM_{2.5}$ were analyzed to understand how $PM_{2.5}$ varied hour by hour. **Figure 1** shows the monthly distribution of AQI categories based on $PM_{2.5}$. Forecasters used this information to help guide their predictions. For example, unhealthy AQI levels were more likely to occur in December and January and less likely to occur in October, November, February, and March. **Figure 2** shows the AQI frequency by day of week and indicates a slight tendency to measure poorer air quality during Thursday through Saturday.

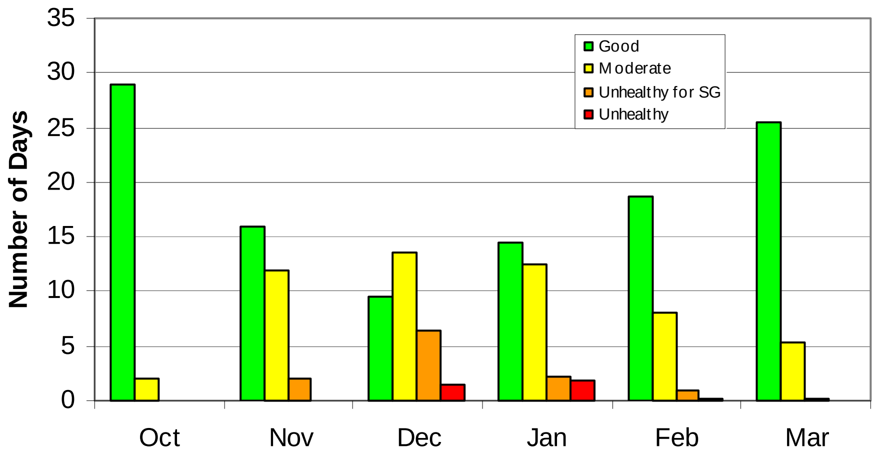


Figure 1. Monthly distribution of the average number of days in each AQI category from 1999 to 2001. AQI category was based on the peak 24-hr average $PM_{2.5}$ concentrations measured in the greater Salt Lake City region.

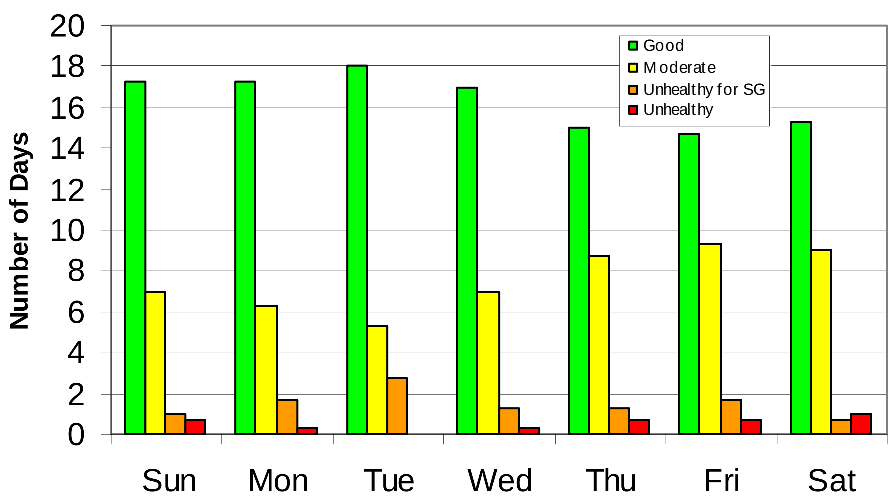


Figure 2. Day-of-week distribution of the average number of days in each AQI category from 1999 to 2001. AQI category was based on the peak 24-hr average $PM_{2.5}$ concentrations measured in the greater Salt Lake City region.

After examining the climatology, physical relationships between meteorological processes and PM were studied to develop a conceptual understanding. We reviewed historical weather maps during PM episodes to understand the large-scale meteorological conditions that produce high PM concentrations. In addition, we performed statistical analyses to relate weather variables to PM concentrations. This conceptual understanding of how meteorology affects PM concentrations was important for two reasons:
(1) forecasters use it to subjectively forecast air quality and (2) understanding the important physical processes guided the selection of potential predictor variables for

developing an objective forecasting tool. These forecasting tools are discussed in Section 3.

4. Daily forecasting. STI issued forecasts seven days a week from January 15 to February 28, 2002. Each day, meteorologists performed the following activities:
   - Verified that current data was received (0900–1000 MST).
   - Reviewed previous day's observations and forecast; reviewed current day's air quality data, weather forecast data, and National Weather Service discussions; ran regression and classification and regression tree (CART) prediction tools; and created current- and next-day forecasts (1100 MST).
   - E-mailed forecast (**Figure 3**) to UDEQ and posted forecasts on a web page as shown in **Figure 4** (1200 MST).
   - Monitored conditions for any dynamic changes (afternoon).

5. Evaluation. Forecast performance was evaluated by comparing forecasted to observed $PM_{2.5}$ levels. This performance evaluation was completed at the conclusion of the project and details are provided in Section 4.

Daily PM2.5 Forecast for Salt Lake City, UT

Today's Date: January 25, 2002

Yesterday's regional maximum AQI:	49 (15 µg/m3) - Good
Today's forecasted regional maximum AQI:	60 (20 µg/m3) - Moderate
Tomorrow's forecasted regional maximum AQI:	60 (20 µg/m3) - Moderate

Discussion:
The upper-level low pressure system that was expected to move towards Utah today has slowed some resulting in a stronger inversion than originally anticipated. The models keep the inversion intact today despite increased winds aloft. This will result in Moderate PM2.5 values for today. The upper-level low pressure system moves towards Utah tomorrow increasing winds aloft; however, the inversion is forecasted to remain intact through Sunday resulting in Moderate PM2.5 levels.

Figure 3. Example e-mail of the daily $PM_{2.5}$ forecast for the forecast issued on January 25, 2002.

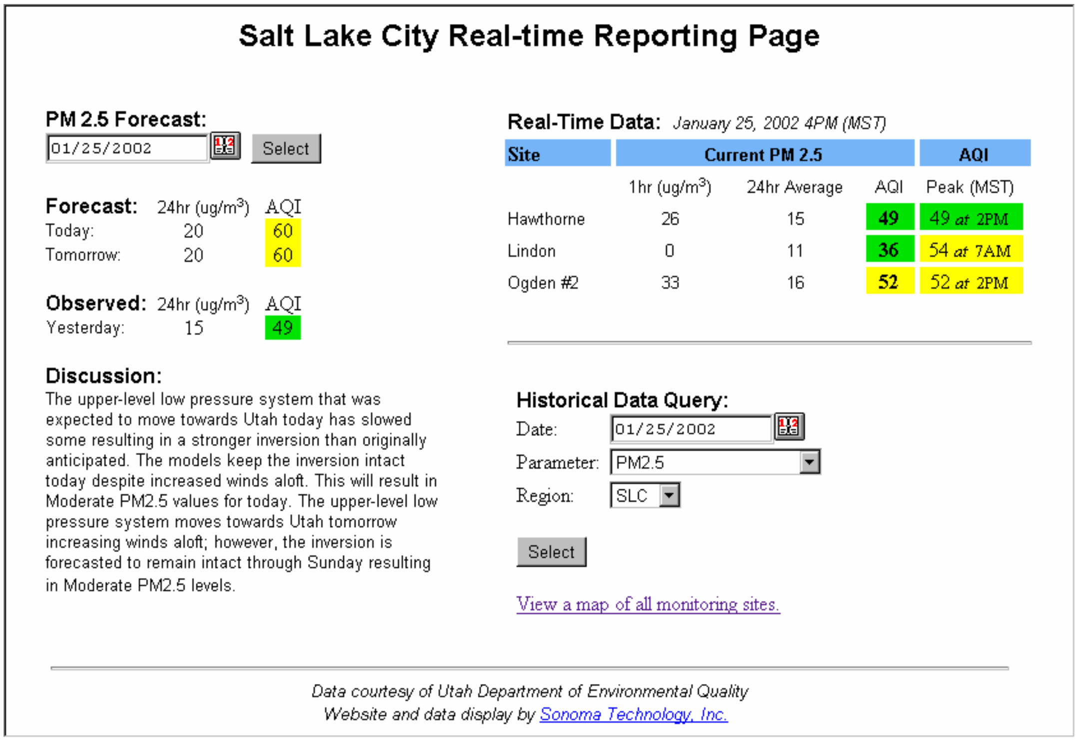


Figure 4. Web page for reviewing the current- and next-day forecasts as well as current conditions.

3. FORECASTING METHODS

Three forecasting methods were developed using a variety of forecasting techniques, both subjective and objective. These include forecasting guidelines, regression equations, and CART, which are discussed in this section. Forecasters can use these tools to aid in predicting daily $PM_{2.5}$ concentrations similar to the way STI forecasters used them during the pilot project.

3.1 FORECASTING GUIDELINES

Forecasting guidelines is a subjective forecasting method that uses specific meteorological criteria to estimate $PM_{2.5}$ concentrations, as shown in **Table 1**. The guidelines were developed using the conceptual model and data analysis results that identified how meteorological features affect $PM_{2.5}$ concentrations in the Salt Lake City region. The column headings indicate the predicted 24-hr average $PM_{2.5}$ concentration range. The typical (or expected) meteorological conditions for the concentration ranges are listed below each column.

Table 1. Forecasting guidelines for the Salt Lake City region.

Guideline	Predicted 24-hr Average $PM_{2.5}$ Concentration Range			
	0-15 μg/m^3	15-25 μg/m^3	25-40 μg/m^3	>40 μg/m^3
	Good	Moderate	Moderate	Unhealthy for Sensitive Groups
Upper-air Pattern	Strong trough overhead	Weak ridge or first day of a moderate ridge overhead or weak trough	Second day of a moderate ridge overhead	Moderate to strong ridge in place for at least two days
Inversion Strength and Duration	No inversion	Weak to moderate inversion that may break	Moderate inversion throughout the day	Moderate to strong inversion persisting throughout the day
Surface Wind Strength		Light to moderate surface winds	Light surface winds	Light surface winds
Surface Pressure Pattern			Weak to moderate surface high over Utah	Moderate to strong surface high overhead
Surface Temperatures			Cool surface temperatures	Cool surface temperatures
Other				Haze or fog in the morning Decoupled surface and aloft winds in the afternoon
Holiday	Add 0 μg/m^3	Add 10 μg/m^3	Add 20 μg/m^3	Add 20 μg/m^3

To use the guidelines, a forecaster needs to first determine the predicted aloft synoptic pattern at 500 mb by finding the general pattern under one of the ranges and continuing to check the other guidelines. If all the forecasted guidelines are met under a given category, the 24-hr $PM_{2.5}$ concentration will be in that range. If any criteria are not met for any given range, use the other forecasting tools described below.

In addition to meteorological variables, a holiday variable was used. Historical analysis showed that $PM_{2.5}$ concentrations increased significantly on holidays, which is likely due to more people staying at home and using their fireplaces. For this reason, major holidays can add as much as 20 ug/m^3 to the 24-hr $PM_{2.5}$ concentrations for the next-day period. Special events such as Super Bowl Sunday can have the same effect, so it is considered a holiday as well. The complete list of holidays includes: Thanksgiving (Thursday through Sunday), Christmas Eve through New Year's Day, Super Bowl Sunday, Veterans' Day, Presidents' Day, Martin Luther King, Jr. Day, Valentine's Day, and weekends immediately following or preceding any of the above days.

In addition to the conditions listed in Table 1, there are several other factors that should be considered, especially if any criteria for a given category are not met.

- Weak short-wave troughs at 700 and 500 mb may not undergo enough vertical mixing to weaken or break the inversion and consequently these features do not lower $PM_{2.5}$ concentrations in the Salt Lake City region.

- The timing of all upper-level trough passages is crucial. The strongest vertical mixing must be directly over Salt Lake City to break the inversion and thereby lower $PM_{2.5}$ concentrations.
- If morning fog or haze are present, $PM_{2.5}$ levels are more likely to be high on that day.
- Decoupled surface and aloft winds in the afternoon are a good indication that the inversion is sufficiently strong.
- $PM_{2.5}$ concentrations rarely increase by more than 20 µg/m^3 per day.

3.2 REGRESSION AND CART FORECASTING METHODS

In addition to subjective forecasting guidelines, objective forecasting methods were also developed. The first, regression, is a statistical method for describing the relationship among variables. In this case, the relationship was between $PM_{2.5}$ concentration and meteorological variables. Two equations were developed, one for a current-day prediction, and one for a next-day prediction. The variables are defined in **Table 2**. The equations are as follows:

Current Day $PM_{2.5}$ = 39.16 + 2.132*Holiday – 0.143*Precip – 0.222*TmaxF
– 0.597*WSres12to00 + 0.689*(700T12Z-TminC)
+ 0.651*(700T00Z-TmaxC) + 0.138*700Td00 – 0.25*850WS00
+ 0.316*YSTPM

Next Day $PM_{2.5}$ = 53.429 + 3.382*Holiday – 0.189*Precip – 0.31*TmaxF
– 0.541*WSres12to00 + 1.008*(700T12Z – TminC)
+ 0.838*(700T00Z – TmaxC) + 0.183*700Td00Z – 0.292*850WS00Z

Table 2. Variables used in the regression equations.

Variable	Description
Holiday	1 for Valentine's Day, Martin Luther King, Jr. Day, Presidents' Day, Veterans' Day, and Super Bowl Sunday. 2 for Thanksgiving weekend and Christmas Eve through New Year's Day. 0 for all other days. 1 for weekends immediately preceding or following any of the above holidays
Precip	Forecasted precipitation in inches during the 24-hr forecast period
TmaxF	Forecasted daytime maximum temperature in ºF
WSres12to00	Average resultant wind speed from 12Z to 00Z (0500 to 1700 MST)
700T12	Temperature at 700 mb at 12Z (0500 MST) in ºC
TminC	Forecasted or observed minimum temperature in ºC
700T00	Temperature at 700 mb at 00Z (1700 MST) in ºC
TmaxC	Forecasted daytime maximum temperature in ºC
700Td00	Dew-point temperature at 700 mb at 00Z (1700 MST) in ºC
850WS00	Wind speed at 850 mb at 00Z (1700 MST) in m/s
YSTPM	Observed 24-hr $PM_{2.5}$ concentration yesterday in µg/m^3

To use the equations, forecasters must acquire the meteorological variables from a reliable source, verify that the values are reasonable, and enter them into the equation. The equation can be used with a simple tool, such as a Microsoft Excel spreadsheet, or with a more advanced tool, such as an operational database that will automatically retrieve the necessary forecast parameters and run the equations.

The second objective forecasting method is CART, a statistical tool that works by classifying data into distinct groups. Like regression equations, CART uses statistical relationships between $PM_{2.5}$ and meteorological variables. CART allows multiple pathways or combinations of variables that will lead to high $PM_{2.5}$ concentrations. The decision tree for next-day $PM_{2.5}$ concentration is shown in **Figure 5** and the variables are defined in **Table 3**.

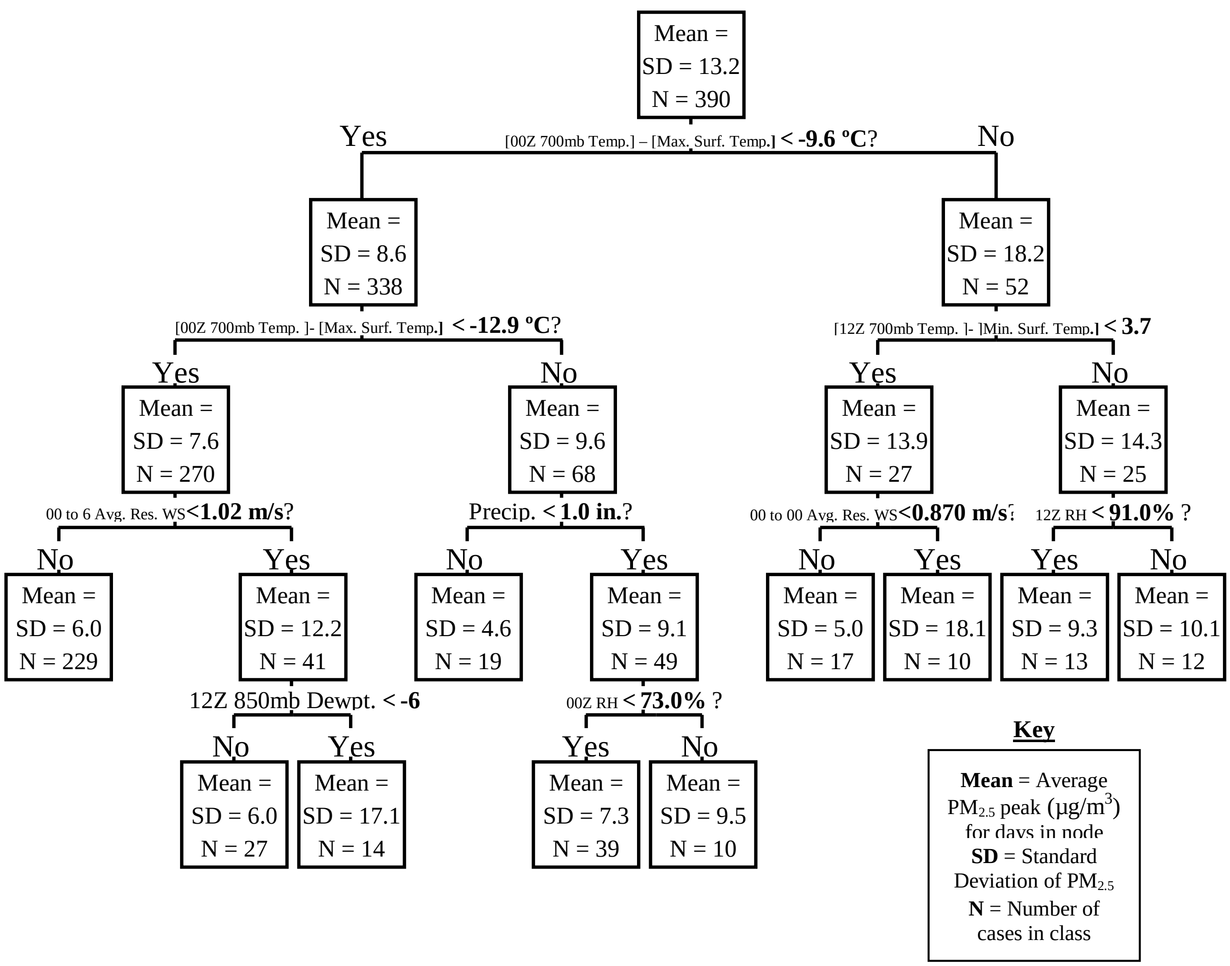


Figure 5. Forecasting decision tree to predict the next-day's peak $PM_{2.5}$ concentrations in the Salt Lake City region.

Table 3. Variables used in CART for Salt Lake City.

Variable	Description
700T00 – TmaxC	Difference between the 700 mb temperature at 00Z (1700 MST) and the daytime maximum temperature in ºC
700T12 – TminC	Difference between the 700 mb temperature at 12Z (0500 MST) and the minimum surface temperature in ºC
WSres12a6a	Average resultant wind speed from midnight to 0600 MST
WSres24	24-hr average resultant wind speed from midnight to midnight MST
Precip	Total precipitation for the 24-hr forecast period
850Td12	Dew-point temperature at 850 mb at 12Z in ºC
12ZRH	Surface relative humidity at 12Z
00ZRH	Surface relative humidity at 00Z

To use CART, forecasters view the criterion at the top of the decision tree and determine whether it will be met. If met, move down the “Yes” side of the tree, if not met, move down the “No” side of the tree. Proceed down the tree until an end node is reached. Each end node has the mean concentration for that group of historical data and thus the predicted $PM_{2.5}$ concentration for the next day.

Note that these forecasting methods are based on the historical data collected by the Federal Reference Method (FRM) monitors. Thus, the methods are “calibrated” to forecast the FRM data. However, when actually predicting PM during the winter of 2001-2002, forecasters did not have access to the FRM data (because it requires laboratory post-analysis) and they relied on the PM data measured from continuous monitors (like the TEOM). We found that the continuous monitors measured $PM_{2.5}$ concentrations that are significantly lower than the reference FRM measurements. This low bias in the continuous data is shown by a scatter plot in **Figure 6**, which compares continuous and FRM $PM_{2.5}$ measurements. Without correcting for this low bias, the forecasts using the above methods might appear to overpredict when verified against the continuous data.

To account for this low bias in the data, UDEQ developed a corrected factor to apply to the continuous data to simulate the FRM data. We used three years of continuous data from collocated TEOM and FRM instruments to develop the correction factor. Linear regression was applied to these data to create the following correction equation:

$$\text{TEOMc} = -1.98 + 1.08\ \text{TEOM} - 0.67\ \text{Avg. Temp.} + 0.08\ \text{Avg. RH}$$

where:
TEOMc = corrected TEOM $PM_{2.5}$ concentration
Avg. Temp = Average 24-hr temperature (°F)
Avg. RH = Average 24-hr relative humidity

By applying this correction to the $PM_{2.5}$ data, we adjusted the continuous data to better match the FRM data as shown in **Figure 7**. This correction was also applied in real-time to adjust the continuous data from the TEOM to better match the historical FRM measurements.

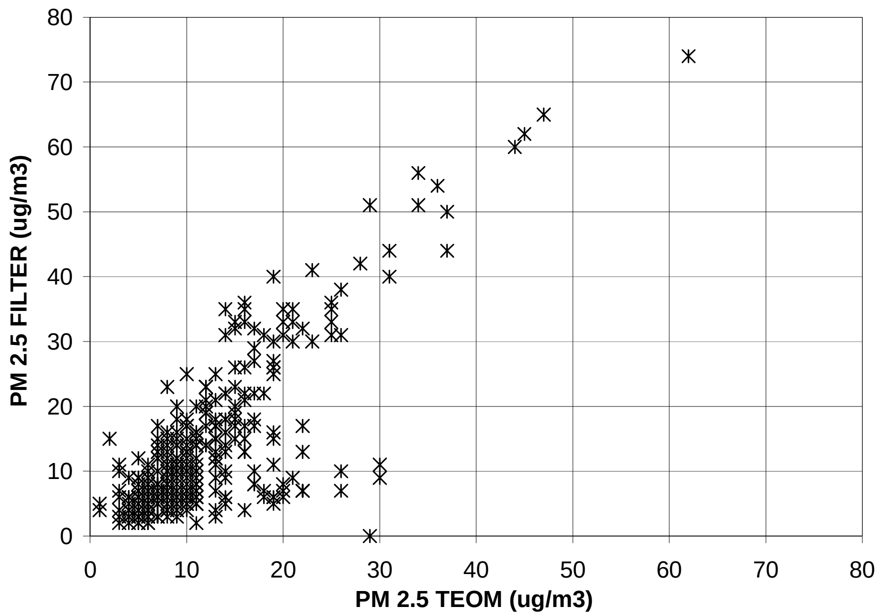


Figure 6. Scatter plot showing the 24-hr average $PM_{2.5}$ measurements collected from collocated FRM and TEOM measurements at the Lindon site.

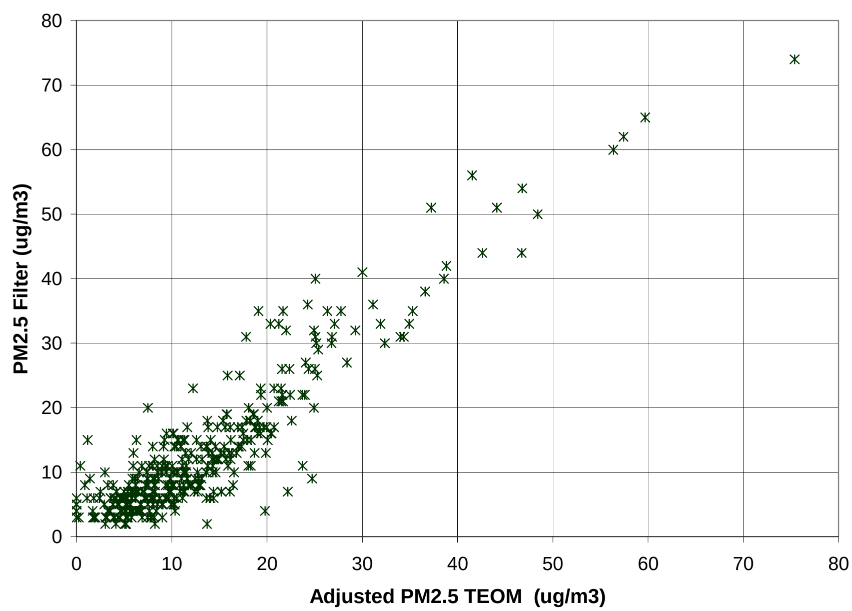


Figure 7. Scatter plot showing the 24-hr average $PM_{2.5}$ measurements collected from collocated FRM and the corrected TEOM measurements at the Lindon site.

4. FORECAST PERFORMANCE

Performance of a forecast can be measured using numerous quantitative and qualitative criteria. For this program, we chose to evaluate how well the forecasts "tracked" with the observations and to examine how the forecasts were used by media. Each day, meteorologists issued a forecast for the peak $PM_{2.5}$ concentrations for the current- and next-day periods. Forecasts are for the maximum 24-hr average $PM_{2.5}$ concentrations from midnight to midnight observed at any of the three continuous $PM_{2.5}$ monitors in the region.

Performance of the current- and next-day forecasts is shown in **Figure 8** for the 45-day forecasting period. This figure illustrates the following:

- Next-day forecasts were not as accurate as current-day forecasts; this is generally expected due to the increased uncertainty of longer-range weather forecasts.
- Both the current- and next-day forecasts accurately predicted the general increasing and decreasing trends in $PM_{2.5}$ concentrations over the 45-day period.
- Forecasters accurately predicted the increasing $PM_{2.5}$ concentrations prior to and at the beginning of the Olympics.
- More volatility in the forecasts and observations occurred, which may partly be due to different (or unusual) emissions patterns during the Olympics (February 8 to 24, 2002).

Forecast performance can also be evaluated antidotally by examining who used the air quality forecasts. First, UDEQ utilized the daily forecasts to help plan their media and outreach strategies during the Olympics, especially to anticipate days with poor air quality. Several other media sources used the $PM_{2.5}$ forecasts. *USA Today* published the air quality forecast several times on the weather page as shown in **Figure 9**. In addition, The Weather Channel showed several graphics of the forecasts and highlighted the air quality conditions during the Olympics.

In general, we consider the forecast performance to be quite good considering several challenging factors: the quick ramp-up time to develop the program (two months), the data issues identified with continuous and FRM measurements, and the likelihood of different emissions patterns due to the Olympics.

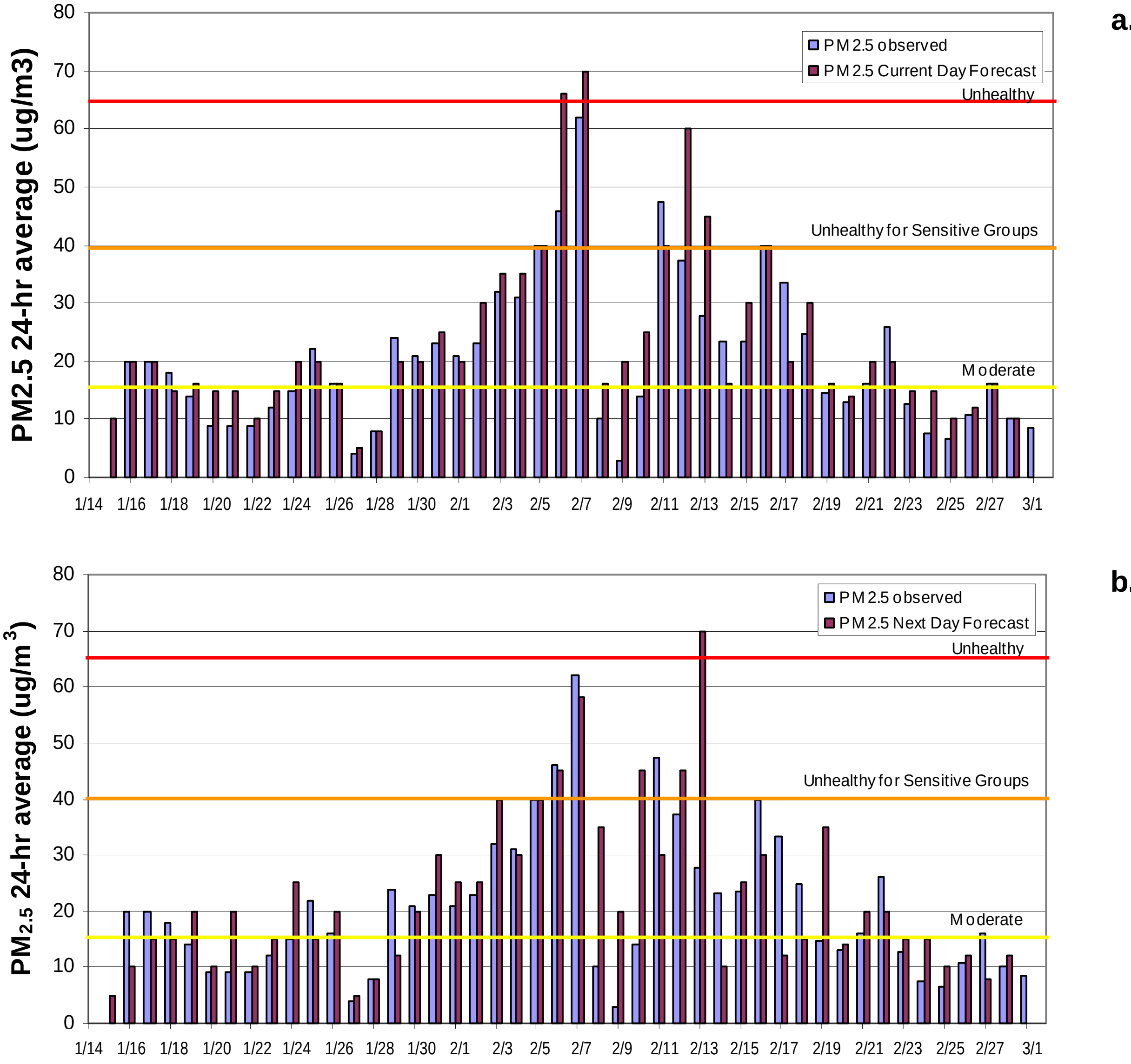
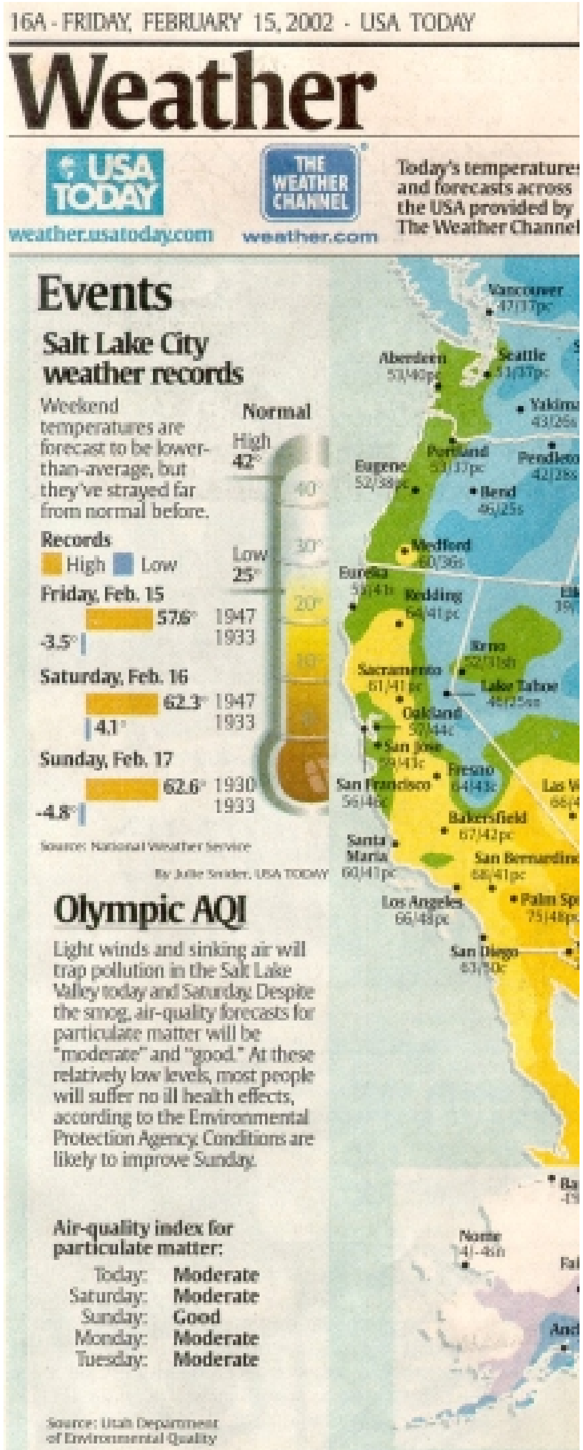


Figure 8. Time series plots of (a) current-day and (b) next-day forecasts issued by meteorologists compared to the observed $PM_{2.5}$ during the 45-day pilot forecasting period.



16A · FRIDAY, FEBRUARY 15, 2002 · USA TODAY

Weather

USA TODAY weather.usatoday.com — THE WEATHER CHANNEL weather.com

Today's temperatures and forecasts across the USA provided by The Weather Channel

Events

Salt Lake City weather records

Weekend temperatures are forecast to be lower-than-average, but they've strayed far from normal before.

Normal: High 42°, Low 25°

Records: High / Low

Friday, Feb. 15: 57.6° 1947; -3.5° 1933

Saturday, Feb. 16: 62.3° 1947; 4.1° 1933

Sunday, Feb. 17: 62.6° 1930; -4.8° 1933

Source: National Weather Service

By Julie Snider, USA TODAY

Olympic AQI

Light winds and sinking air will trap pollution in the Salt Lake Valley today and Saturday. Despite the smog, air-quality forecasts for particulate matter will be "moderate" and "good." At these relatively low levels, most people will suffer no ill health effects, according to the Environmental Protection Agency. Conditions are likely to improve Sunday.

Air-quality index for particulate matter:

Today:	Moderate
Saturday:	Moderate
Sunday:	Good
Monday:	Moderate
Tuesday:	Moderate

Source: Utah Department of Environmental Quality

Figure 9. USA Today article on air quality forecast during the Olympics published on February 15, 2002.

5. RECOMMENDATIONS

We recommend the following to develop and improve the $PM_{2.5}$ forecasting program:

- Further examine ways to adjust the TEOM data to better "simulate" the FRM measurements. Explore ways to use the ambient temperature and humidity data to develop correction factors to apply in real time. In addition, examine the seasonal differences in these two types of measurements.
- Develop a more extensive climatology of $PM_{2.5}$ data and evaluate seasonal, monthly, day of week, and diurnal distribution to better understand the characteristics of $PM_{2.5}$ episodes.
- Conduct case studies to evaluate the local and large-scale weather conditions during severe $PM_{2.5}$ episodes and document this information so it can be used by future forecasters.
- Implement the regression equations and CART decision tree. Test these forecasting tools for several winter seasons. When the tools mis-predict, evaluate what caused the mis-prediction.

6. REFERENCES

Dye T.S., Miller D.S., MacDonald C.P., Anderson C.B., and Thompson B.S. (2002) $PM_{2.5}$ forecasting method development and operations for Salt Lake City, Utah. Presented at *U.S. Environmental Protection Agency's National Air Conference: Forecasting and Public Outreach, San Francisco, CA, February 4-6, 2002*, (STI-2146).

U.S. Environmental Protection Agency (1999) Guideline for developing an ozone forecasting program. Report prepared for Office of Air Quality Planning and Standards Environmental Protection Agency, Research Triangle Park, NC 27711, July 1999, EPA-454/R-99-009 (http://www.epa.gov/ttn/oarpg/t1/memoranda/foreguid.pdf).